



普通高等教育“十一五”国家级规划教材

概率论与数理统计

第四版

□ 浙江大学 盛 骤 谢式千 潘承毅 编



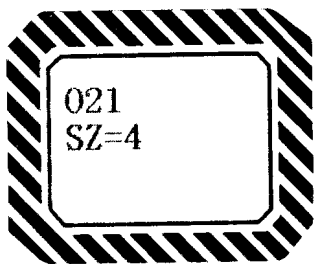
高等教育出版社
HIGHER EDUCATION PRESS

普通高等教育“十一五”国家级规划教材

概率论与数理统计

第四版

浙江大学 盛骤 谢式千 潘承毅 编



高等教育出版社

内容提要

本书是普通高等教育“十一五”国家级规划教材,在2001年出版的《概率论与数理统计》(第三版)的基础上增订而成。本次修订新增的内容有:在数理统计中应用 Excel, bootstrap 方法, p 值检验法, 箱线图;同时吸收了国内外优秀教材的优点对习题的类型和数量进行了调整和充实。

本书主要内容包括概率论、数理统计、随机过程三部分,每章附有习题;同时涵盖了《全国硕士研究生入学统一考试数学考试大纲》的所有知识点。本书可作为高等学校工科、理科(非数学专业)各专业的教材和研究生入学考试的参考书,也可供工程技术人员、科技工作者参考。

图书在版编目(CIP)数据

概率论与数理统计/盛骤,谢式千,潘承毅编. - 4版.
北京:高等教育出版社,2008.6
ISBN 978-7-04-023896-9

I. 概… II. ①盛…②谢…③潘… III. ①概率论-高等学校-教材②数理统计-高等学校-教材 IV. O21

中国版本图书馆 CIP 数据核字(2008)第 057920 号

策划编辑 李 蕊 责任编辑 蒋 青 封面设计 王凌波 责任绘图 杜晓丹
版式设计 余 杨 责任校对 朱惠芳 责任印制 韩 刚

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街4号	免费咨询	800-810-0598
邮政编码	100120	网 址	http://www.hep.edu.cn
总 机	010-58581000		http://www.hep.com.cn
经 销	蓝色畅想图书发行有限公司	网上订购	http://www.landaco.com
印 刷	北京中科印刷有限公司	畅想教育	http://www.widedu.com
开 本	787×960 1/16	版 次	1979年3月第1版
印 张	26.75	印 次	2008年6月第4版
字 数	490 000	定 价	2008年6月第1次印刷 25.80元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 23896-00

第四版前言

本书自1979年3月初版至今,已发行近三十年。历经多年教学实践的检验,得到了国内广大院校和任课教师的认可,发行量为国内同类教材中最多的。

第四版是普通高等教育“十一五”国家级规划教材,在第三版的基础上修订编写而成。在编写之前,高等教育出版社在全国有关高校作过相当广泛的调查,本版的编写吸取了相关的意见。

教材应该力求与时俱进。本版新增加了以下内容:

(1)介绍了bootstrap方法的基本思想和方法,介绍了用bootstrap方法求参数点估计和区间估计的具体做法。bootstrap方法是近代统计中的一种用于数据处理的重要的实用方法。

(2)新增了在数理统计中应用Excel软件一章。介绍了Excel软件及其在数理统计中的应用,举例介绍了应用VBA语言编写“宏”求解具体的数理统计问题。

(3)新增了假设检验问题的 p 值检验法。新增了箱线图,箱线图能大致描述随机变量分布的一些重要性质,还能检测疑似异常点。

(4)对第三版原有的例题和习题作了一些调整,增加了有关加强基本概念、基本运算的习题,在例题和习题的选择上扩大了涉及的范围,例如,农业、保险业、医学、商业、管理学、体育、等等。

选用本教材的院校类别较为广泛,专业不一,学生程度不一。我们认为,教材内容要比教学大纲多一些,要比教师在课堂讲授的多一些,这样能照顾到各类学校各个专业的需要,能满足不同程度的学生的学习需要。

我们在目录中打上了一些*号,在学时限制下,有*号的内容可以不学。这些内容是相对独立的,删去不学不影响全书的讲授。在概率论与数理统计部分中打*号的内容有:基于截尾样本的最大似然估计;置信区间与假设检验之间的关系;样本容量的选取;秩和检验。此外还有偏度、峰度检验,以及这一版新增的部分或全部内容。随机过程部分视教学计划中是否有这一门课决定取舍。

本次修订也包括配套辅导书,它们将与教材同时出版。

本书中新增的有关在数理统计中应用Excel软件的内容由浙江大学于渤教授编写。

本书由浙江大学范大茵教授审阅,对此我们表示衷心的感谢。

高等教育出版社蒋青、李蕊、兰莹莹同志为本版教材做了很多认真、细致的工作,对此,我们表示诚挚的感谢。

诚恳地希望读者批评、指正。

盛 骤 谢式千 潘承毅

2008年4月

第三版前言

这一版我们对于本书第二版中的一些疏漏和不妥之处作了修改,增加了“基于截尾样本的最大似然估计”和“置信区间与假设检验之间的关系”两小节,对各章的例题和习题作了少量的增减。

为了帮助读者抓住要点,提高学习质量与效率,在各章末增写了“小结”。小结中所包含的内容,有的是用来说明概念的现实背景和含义,对某些概念与方法所基于的概率和统计思想作了进一步的阐述;有的则阐明一章内容的重点和基本要求;有的则指出学习时应注意之点。小结也能起到提纲挈领的作用。

书末还增加了两个参读材料:(一)随机变量样本值的产生,(二)标准正态变量分布函数 $\Phi(x)$ 的数值计算。这些内容在解决实际问题时是常会用到的。

本书这一版承柴根象教授、王静龙教授、谢国瑞教授、范大茵教授审阅,他们提出了许多宝贵意见,对此我们表示衷心的感谢。

盛 骤 谢式千 潘承毅

2000年8月

第二版前言

本书是在1979年出版的第一版的基础上修订的,可作为高等学校工科、理科(非数学专业)概率论与数理统计课程的教材,也可供工程技术人员参考。

本书分三部分。概率论部分(第一章至第五章)作为基础知识,为读者提供了必要的理论基础。数理统计部分(第六章至第九章)主要讲述了参数估计和假设检验,并介绍了方差分析和回归分析。随机过程部分(第十章至第十二章)在讲清基本知识的基础上主要讨论了平稳随机过程,还介绍了马尔可夫过程。数理统计和随机过程这两部分内容是相互独立的,可根据专业的需要选用。

在本书第一版出版后,我们经过进一步的教学实践,积累了不少的经验,并吸收了广大读者的意见,修订稿是在这一基础上写出的。我们修改了第一版中存在的不当之处,并致力于教材质量的提高。我们在选材和叙述上尽量做到联系工科专业的实际,注重应用,力图将概念写得清晰易懂,做到便于教学。我们在例题和习题的选择上作了努力,这些题目既具有启发性,又有广泛的应用性,从题目的广泛性也可看到本门课程涉及生活和技术应用领域的广泛性。读者将会发现,这些例题和习题是饶有趣味的。为适应经济建设的需要,我们加强了数理统计的内容,例如编写了“矩估计法”、“样本容量的选取”和“正态分布的偏度、峰度检验”等,并有意识地加强学习统计计算能力的培养。

书中的一部分内容能直接应用于解决实际课题,另一部分内容为读者今后进一步学习有关课程或在实际应用方面提供一定的基础。

黄纪青同志曾参加过本书第一版编写大纲的讨论,撰写过第一版第一章的初稿。

本书的全部插图是由张礼明同志描绘的。

本书第二版承魏宗舒教授、林少官教授、沈恒范教授、范大茵副教授、樊孝述副教授和汪振鹏副教授审阅,他们提出了很多宝贵意见,对此我们表示衷心的感谢。

书中不足之处,诚恳地希望读者批评指正。

盛 骤 谢式千 潘承毅

1988年1月

目 录

第四版前言	I
第三版前言	I
第二版前言	I
第一章 概率论的基本概念	1
§ 1 随机试验	1
§ 2 样本空间、随机事件	2
§ 3 频率与概率	5
§ 4 等可能概型(古典概型)	9
§ 5 条件概率	14
§ 6 独立性	20
小结	23
习题	24
第二章 随机变量及其分布	30
§ 1 随机变量	30
§ 2 离散型随机变量及其分布律	32
§ 3 随机变量的分布函数	38
§ 4 连续型随机变量及其概率密度	42
§ 5 随机变量的函数的分布	50
小结	54
习题	55
第三章 多维随机变量及其分布	60
§ 1 二维随机变量	60
§ 2 边缘分布	64
§ 3 条件分布	67
§ 4 相互独立的随机变量	72
§ 5 两个随机变量的函数的分布	76
小结	83
习题	84
第四章 随机变量的数字特征	90
§ 1 数学期望	90

§ 2 方差	100
§ 3 协方差及相关系数	106
§ 4 矩、协方差矩阵	110
小结	112
习题	113
第五章 大数定律及中心极限定理	119
§ 1 大数定律	119
§ 2 中心极限定理	121
小结	126
习题	126
第六章 样本及抽样分布	128
§ 1 随机样本	128
§ 2 直方图和箱线图	130
§ 3 抽样分布	135
小结	144
附录	145
习题	147
第七章 参数估计	149
§ 1 点估计	149
§ 2 基于截尾样本的最大似然估计	156
§ 3 估计量的评选标准	158
§ 4 区间估计	161
§ 5 正态总体均值与方差的区间估计	163
§ 6 (0-1)分布参数的区间估计	168
§ 7 单侧置信区间	169
小结	170
习题	173
第八章 假设检验	178
§ 1 假设检验	178
§ 2 正态总体均值的假设检验	183
§ 3 正态总体方差的假设检验	187
§ 4 置信区间与假设检验之间的关系	192
§ 5 样本容量的选取	193
§ 6 分布拟合检验	198
§ 7 秩和检验	208

§ 8 假设检验问题的 p 值检验法	213
小结	217
习题	218
第九章 方差分析及回归分析	224
§ 1 单因素试验的方差分析	224
§ 2 双因素试验的方差分析	233
§ 3 一元线性回归	244
§ 4 多元线性回归	257
小结	261
附录	263
习题	265
第十章 bootstrap 方法	270
§ 1 非参数 bootstrap 方法	270
§ 2 参数 bootstrap 方法	278
小结	281
第十一章 在数理统计中应用 Excel 软件	282
§ 1 概述	282
§ 2 箱线图	284
§ 3 假设检验	285
§ 4 方差分析	287
§ 5 一元线性回归	291
§ 6 bootstrap 方法、宏、VBA	293
本章参考文献	299
第十二章 随机过程及其统计描述	300
§ 1 随机过程的概念	300
§ 2 随机过程的统计描述	303
§ 3 泊松过程及维纳过程	309
小结	316
习题	317
第十三章 马尔可夫链	319
§ 1 马尔可夫过程及其概率分布	319
§ 2 多步转移概率的确定	325
§ 3 遍历性	328
小结	331
习题	333

第十四章 平稳随机过程	335
§ 1 平稳随机过程的概念	335
§ 2 各态历经性	338
§ 3 相关函数的性质	346
§ 4 平稳随机过程的功率谱密度	348
小结	358
习题	360
选做习题	363
参读材料 随机变量样本值的产生	376
附表	379
附表 1 几种常用的概率分布表	379
附表 2 标准正态分布表	382
附表 3 泊松分布表	383
附表 4 t 分布表	385
附表 5 χ^2 分布表	386
附表 6 F 分布表	387
附表 7 均值的 t 检验的样本容量	392
附表 8 均值差的 t 检验的样本容量	394
附表 9 秩和临界值表	396
习题答案	397

第一章 概率论的基本概念

自然界和社会上发生的现象是多种多样的. 有一类现象, 在一定条件下必然发生, 例如, 向上抛一石子必然下落, 同性电荷必相互排斥, 等等. 这类现象称为**确定性现象**. 在自然界和社会上存在着另一类现象, 例如, 在相同条件下抛同一枚硬币, 其结果可能是正面朝上, 也可能是反面朝上, 并且在每次抛掷之前无法肯定抛掷的结果是什么; 用同一门炮向同一目标射击, 各次弹着点不尽相同, 在一次射击之前无法预测弹着点的确切位置. 这类现象, 在一定的条件下, 可能出现这样的结果, 也可能出现那样的结果, 而在试验或观察之前不能预知确切的结果. 但人们经过长期实践并深入研究之后, 发现这类现象在大量重复试验或观察下, 它的结果却呈现出某种规律性. 例如, 多次重复抛一枚硬币得到正面朝上大致有一半, 同一门炮射击同一目标的弹着点按照一定规律分布, 等等. 这种在大量重复试验或观察中所呈现出的固有规律性, 就是我们以后所说的**统计规律性**.

这种在个别试验中其结果呈现出不确定性, 在大量重复试验中其结果又具有统计规律性的现象, 我们称之为**随机现象**. 概率论与数理统计是研究和揭示随机现象统计规律性的一门数学学科.

§1 随机试验

我们遇到过各种试验. 在这里, 我们把试验作为一个含义广泛的术语. 它包括各种各样的科学实验, 甚至对某一事物的某一特征的观察也认为是一种试验. 下面举一些试验的例子.

E_1 : 抛一枚硬币, 观察正面 H 、反面 T 出现的情况.

E_2 : 将一枚硬币抛掷三次, 观察正面 H 、反面 T 出现的情况.

E_3 : 将一枚硬币抛掷三次, 观察出现正面的次数.

E_4 : 抛一颗骰子, 观察出现的点数.

E_5 : 记录某城市 120 急救电话台一昼夜接到的呼唤次数.

E_6 : 在一批灯泡中任意抽取一只, 测试它的寿命.

E_7 : 记录某地一昼夜的最高温度和最低温度.

上面举出了七个试验的例子, 它们有着共同的特点. 例如, 试验 E_1 有两种可能结果, 出现 H 或者出现 T , 但在抛掷之前不能确定出现 H 还是出现 T , 这

个试验可以在相同的条件下重复地进行. 又如试验 E_6 , 我们知道灯泡的寿命(以小时计) $t \geq 0$, 但在测试之前不能确定它的寿命有多长. 这一试验也可以在相同的条件下重复地进行. 概括起来, 这些试验具有以下的特点:

1° 可以在相同的条件下重复地进行;

2° 每次试验的可能结果不止一个, 并且能事先明确试验的所有可能结果;

3° 进行一次试验之前不能确定哪一个结果会出现.

在概率论中, 我们将具有上述三个特点的试验称为**随机试验**.

本书中以后提到的试验都是指随机试验.

我们是通过研究随机试验来研究随机现象的.

§ 2 样本空间、随机事件

(一) 样本空间

对于随机试验, 尽管在每次试验之前不能预知试验的结果, 但试验的所有可能结果组成的集合是已知的. 我们将随机试验 E 的所有可能结果组成的集合称为 E 的**样本空间**, 记为 S . 样本空间的元素, 即 E 的每个结果, 称为**样本点**.

下面写出 § 1 中试验 E_k ($k=1, 2, \dots, 7$) 的样本空间 S_k :

$$S_1: \{H, T\};$$

$$S_2: \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\};$$

$$S_3: \{0, 1, 2, 3\};$$

$$S_4: \{1, 2, 3, 4, 5, 6\};$$

$$S_5: \{0, 1, 2, 3, \dots\};$$

$$S_6: \{t | t \geq 0\};$$

$S_7: \{(x, y) | T_0 \leq x \leq y \leq T_1\}$, 这里 x 表示最低温度($^{\circ}\text{C}$), y 表示最高温度($^{\circ}\text{C}$). 并设这一地区的温度不会小于 T_0 , 也不会大于 T_1 .

(二) 随机事件

在实际中, 当进行随机试验时, 人们常常关心满足某种条件的那些样本点所组成的集合. 例如, 若规定某种灯泡的寿命(小时)小于 500 为次品, 则在 E_6 中我们关心灯泡的寿命是否有 $t \geq 500$. 满足这一条件的样本点组成 S_6 的一个子集: $A = \{t | t \geq 500\}$. 我们称 A 为试验 E_6 的一个**随机事件**. 显然, 当且仅当子集 A 中的一个样本点出现时, 有 $t \geq 500$.

一般,我们称试验 E 的样本空间 S 的子集为 E 的随机事件,简称事件^①. 在每次试验中,当且仅当这一子集中的一个样本点出现时,称这一事件发生.

特别,由一个样本点组成的单点集,称为基本事件. 例如,试验 E_1 有两个基本事件 $\{H\}$ 和 $\{T\}$; 试验 E_4 有 6 个基本事件 $\{1\}, \{2\}, \dots, \{6\}$.

样本空间 S 包含所有的样本点,它是 S 自身的子集,在每次试验中它总是发生的, S 称为必然事件. 空集 \emptyset 不包含任何样本点,它也作为样本空间的子集,它在每次试验中都不发生, \emptyset 称为不可能事件.

下面举几个事件的例子.

例 1 在 E_2 中事件 A_1 : “第一次出现的是 H ”, 即

$$A_1 = \{HHH, HHT, HTH, HTT\}.$$

事件 A_2 : “三次出现同一面”, 即

$$A_2 = \{HHH, TTT\}.$$

在 E_6 中, 事件 A_3 : “寿命小于 1000 小时”, 即

$$A_3 = \{t | 0 \leq t < 1000\}.$$

在 E_7 中, 事件 A_4 : “最高温度与最低温度相差 10 摄氏度”, 即

$$A_4 = \{(x, y) | y - x = 10, T_0 \leq x \leq y \leq T_1\}.$$

□

(三) 事件间的关系与事件的运算

事件是一个集合, 因而事件间的关系与事件的运算自然按照集合论中集合之间的关系和集合运算来处理. 下面给出这些关系和运算在概率论中的提法. 并根据“事件发生”的含义, 给出它们在概率论中的含义.

设试验 E 的样本空间为 S , 而 $A, B, A_k (k=1, 2, \dots)$ 是 S 的子集.

1° 若 $A \subset B$, 则称事件 B 包含事件 A , 这指的是事件 A 发生必导致事件 B 发生.

若 $A \subset B$ 且 $B \subset A$, 即 $A = B$, 则称事件 A 与事件 B 相等.

2° 事件 $A \cup B = \{x | x \in A \text{ 或 } x \in B\}$ 称为事件 A 与事件 B 的和事件. 当且仅当 A, B 中至少有一个发生时, 事件 $A \cup B$ 发生.

类似地, 称 $\bigcup_{k=1}^n A_k$ 为 n 个事件 A_1, A_2, \dots, A_n 的和事件; 称 $\bigcup_{k=1}^{\infty} A_k$ 为可列个事件 A_1, A_2, \dots 的和事件.

3° 事件 $A \cap B = \{x | x \in A \text{ 且 } x \in B\}$ 称为事件 A 与事件 B 的积事件. 当且仅

^① 严格地说, 事件是指 S 中的满足某些条件的子集. 当 S 是由有限个元素或由可列无限个元素组成时, 每个子集都可作为一个事件. 若 S 是由不可列无限个元素组成时, 某些子集必须排除在外. 幸而这种不可容许的子集在实际应用中几乎不会遇到. 今后, 每当我们讲到一个事件时都是假定它是容许考虑的那种子集. 读者如有兴趣可参考较详细的教材.

当 A, B 同时发生时, 事件 $A \cap B$ 发生. $A \cap B$ 也记作 AB .

类似地, 称 $\bigcap_{k=1}^n A_k$ 为 n 个事件 A_1, A_2, \dots, A_n 的积事件; 称 $\bigcap_{k=1}^{\infty} A_k$ 为可列个事件 A_1, A_2, \dots 的积事件.

4° 事件 $A - B = \{x | x \in A \text{ 且 } x \notin B\}$ 称为事件 A 与事件 B 的差事件. 当且仅当 A 发生、 B 不发生时事件 $A - B$ 发生.

5° 若 $A \cap B = \emptyset$, 则称事件 A 与 B 是互不相容的, 或互斥的. 这指的是事件 A 与事件 B 不能同时发生. 基本事件是两两互不相容的.

6° 若 $A \cup B = S$ 且 $A \cap B = \emptyset$, 则称事件 A 与事件 B 互为逆事件. 又称事件 A 与事件 B 互为对立事件. 这指的是对每次试验而言, 事件 A, B 中必有一个发生, 且仅有一个发生. A 的对立事件记为 \bar{A} . $\bar{A} = S - A$.

用图 1-1—图 1-6 可直观地表示以上事件之间的关系与运算. 例如, 在图 1-1 中长方形表示样本空间 S , 圆 A 与圆 B 分别表示事件 A 与事件 B , 事件 B 包含事件 A . 又如在图 1-2 中长方形表示样本空间 S , 圆 A 与圆 B 分别表示事件 A 与事件 B , 而阴影部分表示和事件 $A \cup B$.

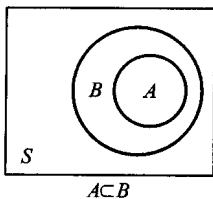


图 1-1

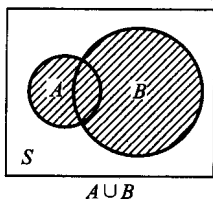


图 1-2

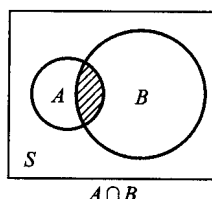


图 1-3

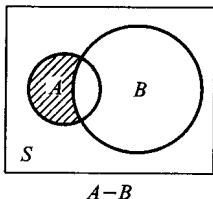


图 1-4

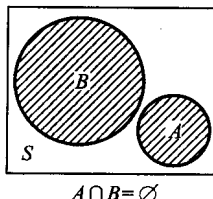


图 1-5

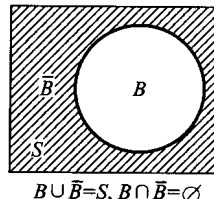


图 1-6

在进行事件运算时, 经常要用到下述定律. 设 A, B, C 为事件, 则有

交换律: $A \cup B = B \cup A$; $A \cap B = B \cap A$.

结合律: $A \cup (B \cap C) = (A \cup B) \cap C$;

$A \cap (B \cup C) = (A \cap B) \cup C$.

分配律: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

德摩根律: $\overline{A \cup B} = \overline{A} \cap \overline{B}$; $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

例 2 在例 1 中有

$$A_1 \cup A_2 = \{HHH, HHT, HTH, HTT, TTT\},$$

$$A_1 \cap A_2 = \{HHH\},$$

$$A_2 - A_1 = \{TTT\},$$

$$\overline{A_1 \cup A_2} = \{THT, TTH, THH\}.$$

□

例 3 如图 1-7 所示的电路中,以 A 表示“信号灯亮”这一事件,以 B, C, D 分别表示事件:继电器接点 I, II, III 闭合,那么容易知道 $BC \subset A, BD \subset A, BC \cup BD = A$, 而 $\overline{BA} = \emptyset$, 即事件 \overline{B} 与事件 A 互不相容. 又, $\overline{B \cup C} = \overline{B} \cap \overline{C}$. (左边表示事件“I, II 至少有一个闭合”的逆事件,也就是 I, II 都不闭合,即 $\overline{B}, \overline{C}$ 同时发生.) □

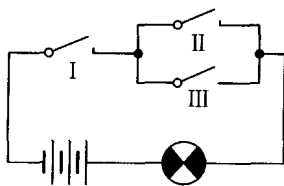


图 1-7

§ 3 频率与概率

对于一个事件(除必然事件和不可能事件外)来说,它在一次试验中可能发生,也可能不发生. 我们常常希望知道某些事件在一次试验中发生的可能性究竟有多大. 例如,为了确定水坝的高度,就要知道河流在造水坝地段每年最大洪水达到某一高度这一事件发生的可能性大小. 我们希望找到一个合适的数来表征事件在一次试验中发生的可能性大小. 为此,首先引入频率,它描述了事件发生的频繁程度,进而引出表征事件在一次试验中发生的可能性大小的数——概率.

(一) 频率

定义 在相同的条件下,进行了 n 次试验,在这 n 次试验中,事件 A 发生的次数 n_A 称为事件 A 发生的频数. 比值 n_A/n 称为事件 A 发生的频率,并记成 $f_n(A)$.

由定义,易见频率具有下述基本性质:

$$1^\circ 0 \leq f_n(A) \leq 1;$$

$$2^\circ f_n(S) = 1;$$

3° 若 A_1, A_2, \dots, A_k 是两两互不相容的事件,则

$$f_n(A_1 \cup A_2 \cup \dots \cup A_k) = f_n(A_1) + f_n(A_2) + \dots + f_n(A_k).$$

由于事件 A 发生的频率是它发生的次数与试验次数之比,其大小表示 A 发

生的频繁程度. 频率大, 事件 A 发生就频繁, 这意味着事件 A 在一次试验中发生的可能性就大. 反之亦然. 因而, 直观的想法是用频率来表示事件 A 在一次试验中发生的可能性的. 但是否可行, 先看下面的例子.

例 1 考虑“抛硬币”这个试验, 我们将一枚硬币抛掷 5 次、50 次、500 次, 各做 10 遍. 得到数据如表 1-1 所示(其中 n_H 表示 H 发生的频数, $f_n(H)$ 表示 H 发生的频率).

表 1-1

实验序号	$n=5$		$n=50$		$n=500$	
	n_H	$f_n(H)$	n_H	$f_n(H)$	n_H	$f_n(H)$
1	2	0.4	22	0.44	251	0.502
2	3	0.6	25	0.50	249	0.498
3	1	0.2	21	0.42	256	0.512
4	5	1.0	25	0.50	253	0.506
5	1	0.2	24	0.48	251	0.502
6	2	0.4	21	0.42	246	0.492
7	4	0.8	18	0.36	244	0.488
8	2	0.4	24	0.48	258	0.516
9	3	0.6	27	0.54	262	0.524
10	3	0.6	31	0.62	247	0.494

这种试验历史上有人做过, 得到如表 1-2 所示的数据.

表 1-2

实验者	n	n_H	$f_n(H)$
德摩根	2 048	1 061	0.518 1
蒲 丰	4 040	2 048	0.506 9
K·皮尔逊	12 000	6 019	0.501 6
K·皮尔逊	24 000	12 012	0.500 5

从上述数据可以看出: 抛硬币次数 n 较小时, 频率 $f_n(H)$ 在 0 与 1 之间随机波动, 其幅度较大, 但随着 n 增大, 频率 $f_n(H)$ 呈现出稳定性. 即当 n 逐渐增大时 $f_n(H)$ 总是在 0.5 附近摆动, 而逐渐稳定于 0.5. \square

例 2 考察英语中特定字母出现的频率. 当观察字母的个数 n (试验的次